Response to Comments on the Manuscript (acp-2020-1004):

“Estimating daily full-coverage and high-accuracy 5-km ambient particulate matters across China: considering their precursors and chemical compositions”

Response to Comments of Referee #2:

General comment:

In this paper, Wang et al. proposed a framework to estimate the daily PM2.5 and PM10 concentration over China by combining multiple data sources with a Light Gradient Boosting Machine learning method. They included satellite product (TROPOMI) and modeled data assimilation dataset (GEOS-FP) as main predictors. Even though they showed some reasonable statistics from the validation process, the advantages of this method are not well justified. Also, some flaws are found in their validation process. As a result, I do not recommend this paper for publication in ACP.

Response: We would like to express our sincere gratitude to the referee for his/her comments and recommendations for improving the paper. An item-by-item response to the comments raised by the referee follows. Thanks for your time.

Major comments:

Q2.1: The selection of predicting variables from TROPOMI is too arbitrary and lacks justification. In this paper, the authors only considered column NOx and SO2 observations, which are the precursors of sulfate and nitrate. Both of them are large components of PM2.5. However, other components are also important. For example organic aerosols. Why the precursors of organic aerosols were not chosen as input predictors? Also, this idea of “PM is
associated with ozone, so choose column ozone as one of the predictors” needs more justifications.

Response: Thank the referee for his/her significant comments. The purpose of our study is to estimate daily full-coverage PM at high spatial resolution using the datasets of their precursors & chemical compositions instead of AOD products. Therefore, the selection of predicting variables should be carefully considered. As for remote sensing sensors, only TROPOMI can generate daily high-spatial-resolution (e.g., 5-km) and high-coverage chemical species at present. In our study, the atmospheric products of NO₂ and SO₂ from TROPOMI were adopted, regarded as two precursors of PM. However, there is merely one precursor of organic aerosol, i.e., formaldehyde, belonging to the atmospheric products from TROPOMI (see https://earth.esa.int/web/guest/missions/esa-eo-missions/sentinel-5p). Since formaldehyde is generally not a major organic aerosol precursor (Hallquist et al., 2009; Volkamer et al., 2006), this product was not utilized. By contrast, the Organic Carbon Column Mass Density (carbon-related) from GEOS-FP was used in this paper. Furthermore, some other chemical compositions of PM were also acquired from GEOS-FP, including the Nitrate Column Mass Density (nitrate-related), SO₄ Column Mass Density (sulfate-related), Black Carbon Column Mass Density (carbon-related), Dust Column Mass Density (dust-related), Ammonium Column Mass Density (ammonium-related), and Sea Salt Column Mass Density (sea salt-related). All of these variables are employed according to the major chemical compositions of PM (i.e., nitrate, sulfate, carbon, dust, ammonium, and sea salt) (Baker et al., 2007; Tucker et al., 2000; Zheng et al., 2005; Wang et al., 2019; Pui et al., 2014). We have appended this statement in the manuscript. It is concluded that the selected variables are sufficient to estimate PM over China. The validation results show that the estimation model achieves a satisfactory performance (e.g., space-based CV R²: 0.88 for PM₂.₅ and 0.83 for PM₁₀) in our study, which also signify this point.

The justification for the adoption of total O₃ column is presented as follows. With regard to stratospheric O₃, a latest study (Chen et al., 2020) has shown that the downward transport of O₃ stemming from the stratosphere-to-troposphere exchange can be a significant contributor to background O₃. Such enhancement of background O₃ will affect ambient PM. In addition, ambient O₃ pollution is rapidly increasing over China in recent years (Liu et al., 2020; Wang et al., 2020) and the proportion of it may also rise in the total O₃ column. At present, the total O₃ column has been used to estimate ambient O₃ over China (Liu et al., 2020) and Tibetan Plateau (Li et al., 2020), suggesting its
surface predictive capacity. In China, ambient PM is associated with ambient O$_3$ (Chen et al., 2019). Therefore, the total O$_3$ column is introduced as an auxiliary variable in our study.

**The main revision is as follows:**

In our study, all of the chemical species are selected in accordance with the major component of PM$_{2.5}$ and PM$_{10}$ (i.e., nitrate, sulfate, carbon, dust, ammonium, and sea salt) (Baker et al., 2007; Tucker et al., 2000; Zheng et al., 2005; Wang et al., 2019; Pui et al., 2014).

**References:**


Q2.2: By using GEOS-FP, this method loses its flexibility to adjust its input conditions. GEOS-FP is one of the data assimilation products from GMAO. It is based largely on the model output of GEOS. The method in this paper is largely based on this dataset. According to their fig 6, 4 out of 6 top feature importance for their PM2.5 prediction are from GEOS-FP. What this means is that their prediction is mostly controlled by a dataset that they could not make any modifications to. For example, the GEOS-FP system has been updated to version 5.25 in January 2020. Let alone some discontinuous issues posed by all the updates along with the release of each GEOS-FP products. On the other hand, the authors argue that previous studies using CTMs have the limitation of large uncertainty in the emission inventories of the CTMs. By saying that, GEOS-FP also has the same problem of large uncertainty in their emission inventories. By using CTMs instead of data assimilation datasets as prediction inputs, researchers can do their best to narrow the uncertainty in their model simulations and could also conduct sensitivity experiments. From this point of view, methods using CTMs or earth models (e.g. GEOS) would be better for this kind of prediction.

Response: Thank the referee for his/her valuable comments. To be clearer, the total comments are divided into two parts, which are replied as follows.

“By using GEOS-FP, this method loses its flexibility to adjust its input conditions. GEOS-FP is one of the data assimilation products from GMAO. It is based largely on the model output of GEOS. The method in this paper is largely based on this dataset. According to their fig 6, 4 out of 6 top feature importance for their PM2.5 prediction are from GEOS-FP. What this means is that their prediction is mostly controlled by a dataset that they could not make any modifications to. For example, the GEOS-FP system has been updated to version 5.25 in January 2020. Let alone some discontinuous issues posed by all the updates along with the release of each GEOS-FP products.”
Figure r1: Daily (20190101) spatial distribution of the GEOS-FP, estimated, and ground-based PM$_{2.5}$. The circles represent the ground-based sites.

Apart from GEOS-FP, multiple datasets from other sources were adopted in our study, including
TROPOMI, MODIS, GPW, and OpenStreetMap. Especially, high-resolution geographical factors, such as NDVI, road density, and population density, could maintain the spatial information. In addition, an advance ensemble learning method, i.e., light gradient boosting machine, was exploited to fuse the multisource data using ground-truth values. Space-based CV results show that the proposed framework performs well in the study area (e.g., $R^2$: 0.88 for PM$_{2.5}$ and 0.83 for PM$_{10}$), suggesting that GEOS-FP data did not introduce large uncertainties. An example to compare the spatial distribution between the estimated and GEOS-FP PM$_{2.5}$ is demonstrated in Figure r1. The GEOS-FP PM$_{2.5}$ is calculated via this formula: PM$_{2.5}$=1.375*SO$_4$+2.1*OC+BC+DS$_{2.5}$+SS$_{2.5}$ (Xiao et al., 2018). As can be seen, the spatial patterns of the estimated PM$_{2.5}$ are much closer to the actual measurements than GEOS-FP, with a higher spatial resolution. Therefore, the estimation accuracy is greatly improved compared to GEOS-FP based on posterior techniques.

“On the other hand, the authors argue that previous studies using CTMs have the limitation of large uncertainty in the emission inventories of the CTMs. By saying that, GEOS-FP also has the same problem of large uncertainty in their emission inventories. By using CTMs instead of data assimilation datasets as prediction inputs, researchers can do their best to narrow the uncertainty in their model simulations and could also conduct sensitivity experiments. From this point of view, methods using CTMs or earth models (e.g. GEOS) would be better for this kind of prediction.”

In this paper, the previous studies using CTMs refer to the works merely considering CTMs without other techniques. Our study does not focus on comparing to or arguing about the methods using CTMs. The mention of CTMs in the introduction was only to elicit that the approaches based on remote sensing satellites have been greatly developed in recent years. The statements about CTMs is confusing and has been rephrased in the manuscript. CTMs and GEOS-FP both potential present large uncertainty in their emission inventories. At present, the outputs of CTMs have been combined with other datasets (e.g., remote sensing) to estimate PM$_{2.5}$, such as the mentioned work (van Donkelaar et al., 2019). In our study, we also fused multiple datasets from new sources, including new remote sensing sensor (TROPOMI) and new data assimilation (GEOS-FP). Moreover, the present study is a novel attempt to estimate daily full-coverage PM based on the datasets of their precursors & chemical compositions instead of AOD products. The validation results show that the proposed framework can perform well
without the input of AOD in the study area.

**The main revision is as follows:**

With regard to CTMs, the uncertainties of the emission inventories could be large in some areas (Li et al., 2017b) and it will consume time and energy to collect the necessary information for simulation (Chu et al., 2016). The approaches based on remote sensing satellites have been greatly developed in recent years (Sorek-Hamer et al., 2020).

**References:**


**Q2.3:** The validation process exists flows. First, the authors defined their “AOD-based” control estimate, which is using VIIRS AOD to replace the TROPOMI and GEOS-FP in their estimating framework. However, this is not the case of previous AOD based studies, especially those most influential ones, for example, van Donkelaar et al. 2019. Previous AOD based studies usually combined with model simulations and ground-based measurements to best use the information of satellite AOD products. What has been done in the paper was that comparing an estimate from a missing data AOD product with an estimate from the combination of satellite observations and model output. So that it caused the issue of comparing apple with orange; moreover, the selling point of this paper is the daily PM2.5 estimate. So validation regarded to a daily resolution would be the most convincing. Otherwise, why not just estimate the seasonal or annual PM concentrations, which are more useful in current epidemiological studies. The validation results from the time-based cross-validation method were the worst according to their table1.

**Response:** Thank the referee for his/her important comments. To fully validate the estimated results, a total of three CV schemes (i.e., sample-based, space-based, and time-based) were considered in our study. Meanwhile, all of the validation results were performed at daily resolution, including the overall validation (Figure 4), regional validation (Table 1), seasonal validation (Fig. S3-S6), and grid-based validation (Figure 5, S7-S9). The examples of daily estimated results in different seasons (20190122,
20190501, 20190803, and 20191111) were also provided in Figure 7 to show the daily spatial distributions.

Moreover, a baseline was devised in our study. Since a large number of related studies over China based on machine learning methods only adopted remote sensing AOD products (no model simulations) to estimate PM (Li et al., 2020; Yang et al., 2019; Wei et al., 2019; He et al., 2018, 2020; Yao et al., 2019; Ma et al., 2016; Chen et al., 2018a, 2018b, 2019; Wang et al., 2019; Zhang et al., 2019; Xue et al., 2020), the VIIRS DB AOD product was selected as the AOD-based. However, related studies could utilize various techniques to improve their estimation performance. For instance, Li et al. (2020) proposed GTWNN; Wei et al. (2019) developed STRF; Kong et al. (2020) combined model simulations and ground-based measurements. It is very difficult to duplicate their experiments as baselines due to different hardware facilities, large time consumptions, closed data sources, unspecified model parameters, etc. By contrast, we compared our metrics to those of related studies over China in recent years, which are listed in Table 1 (or see Table S5 in the supplementary materials). This is a common strategy used in previous studies for comparison with other works (Wei et al., 2019; Jiang et al., 2020; Kong et al., 2020). Apart from Kong et al. (2020) (model simulations), another two related studies over China using model simulations or reanalysis datasets have been appended in the table (Xue et al., 2017; Xiao et al., 2018), as the referee pointed out. The study area of the mentioned work (van Donkelaar et al., 2019) is North America and its temporal resolution is annual. Therefore, this work cannot be compared in our study.

Table 1: Detailed information about the previous related works over China. SACV: sample-based CV; SPCV: space-based CV; TICV: time-based CV; SR: spatial resolution; TR: temporal resolution; FC: full-coverage; T: true; F: false; MF: the factors which lead to the missing values in the estimated results.

<table>
<thead>
<tr>
<th>Type</th>
<th>Reference</th>
<th>Metric</th>
<th>SACV</th>
<th>SPCV</th>
<th>TICV</th>
<th>SR</th>
<th>TR</th>
<th>Study period</th>
<th>FC</th>
<th>MF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td></td>
<td>$R^2$</td>
<td>0.93</td>
<td>0.88</td>
<td>0.73</td>
<td></td>
<td></td>
<td>Daily</td>
<td>2019※</td>
<td>T               None</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>8.87</td>
<td>11.56</td>
<td>17.3</td>
<td>5-</td>
<td>km</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RPE</td>
<td>22.8%</td>
<td>29.8%</td>
<td>44.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R^2$</td>
<td>0.85</td>
<td>0.83</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM$_{2.5}$</td>
<td>Wei et al., 2019</td>
<td>RMSE</td>
<td>15.57</td>
<td>16.63</td>
<td>24.83</td>
<td>1-</td>
<td>km</td>
<td>Daily</td>
<td>2016</td>
<td>F               Cloud, snow/ice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RPE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R^2$</td>
<td>0.8</td>
<td>-</td>
<td>-</td>
<td>3-</td>
<td>km</td>
<td>Daily</td>
<td>2015</td>
<td>F               Cloud, snow/ice,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>18</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>bright surface</td>
</tr>
<tr>
<td>Study</td>
<td>R²</td>
<td>RMSE (μg/m³)</td>
<td>RPE</td>
<td>10-km</td>
<td>Daily</td>
<td>Year</td>
<td>F</td>
<td>Cloud, snow/ice, bright surface</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>------</td>
<td>--------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>---</td>
<td>--------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yao et al., 2019</td>
<td>0.6</td>
<td>21.76</td>
<td></td>
<td></td>
<td></td>
<td>2014</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li et al., 2020</td>
<td>0.79</td>
<td>17.38</td>
<td>31.5%</td>
<td>10-km</td>
<td>Daily</td>
<td>2015</td>
<td>F</td>
<td>Cloud, snow/ice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jiang et al., 2020</td>
<td>0.74</td>
<td>11.02</td>
<td></td>
<td></td>
<td></td>
<td>2018</td>
<td>2018.03.01-2019.02.28 T</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kong et al., 2020</td>
<td>0.86</td>
<td>15.1</td>
<td></td>
<td></td>
<td></td>
<td>2013-2018</td>
<td>T</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xue et al., 2017</td>
<td>0.72</td>
<td>23</td>
<td>41%</td>
<td></td>
<td></td>
<td>2014</td>
<td>T</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xiao et al., 2018</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>2013-2016</td>
<td>T</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposed PM10</td>
<td>0.67</td>
<td>16.92</td>
<td>24.5%</td>
<td></td>
<td></td>
<td>2019</td>
<td>T</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen et al., 2018b</td>
<td>0.45</td>
<td>31.54</td>
<td></td>
<td></td>
<td></td>
<td>2005-2016</td>
<td>F</td>
<td>Cloud, snow/ice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kong et al., 2020</td>
<td></td>
<td>28.8</td>
<td></td>
<td></td>
<td></td>
<td>2013-2018</td>
<td>T</td>
<td>None</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:
1. The symbols of * represent that the works could provide the estimated results at various temporal resolutions, while the metrics listed in the table are computed from the daily estimation.
2. ※: Only the metrics computed from the estimated results through the proposed approach for a whole year (2019) are listed in the table to be fairly compared to previous works. The study period of this paper is from June 1, 2018 to March 31, 2020.

The validation results from the time-based CV were the worst since the temporal heterogeneity of PM is usually strong (Li et al., 2017, 2020). In other word, the temporal variations of PM could not be fully captured. This phenomenon can be discovered in previous related studies over China, such as
Wei et al. (2019) (time-based CV $R^2$: 0.63) and Xiao et al. (2018) (time-based CV $R^2$: 0.73). Compared to them, the time-based CV results in our study are acceptable.

In conclusion, it is believed that the validation process was justified in this paper.

References:


Specific comments:

Q2.4: Line 99: is it necessary to mention that the area of this study was chosen because that China has the largest population in the world? Maybe rephrase this sentence or delete it.

Response: Thank the referee for his/her comment. This statement has been removed in the manuscript.

Q2.5: Line 106: “monitor”? maybe use “estimate”.

Response: Thank the referee for his/her comment. The word “monitor” has been replaced with “estimate” in the manuscript.

The main revision is as follows:

It is necessary to develop an approach that can estimate PM$_{2.5}$ and PM$_{10}$ across China continuously and precisely.

Q2.6: Line 120: any publications about the data validation, calibration, and uncertainty analysis of the Chinese PM2.5/PM10 measurements from CNEMC?

Response: Thank the referee for his/her comment. The required contents have been appended in the manuscript. The ground-based measurements from CNEMC have been widely used to estimate PM across China (Li et al., 2020; Yang et al., 2019; Wei et al., 2019; He et al., 2018, 2020; Yao et al., 2019; Ma et al., 2016; Chen et al., 2018a, 2018b, 2019; Wang et al., 2019; Zhang et al., 2019; Xue et al.,
2017, 2019, 2020; Kong et al., 2020), suggesting their reliability and accuracy.

The main revision is as follows:
The CNEMC can provide hourly ambient concentrations of $\text{PM}_{2.5}$ and $\text{PM}_{10}$ over China, which are obtained according to the technical specification of HJ 817-2018 (i.e., tapered element oscillating microbalance method or beta-attenuation method).

References:


sensing and air quality model simulated data to estimate spatiotemporal variations of PM2.5 exposure in China. Remote Sensing, 9(3), 221.


Q2.7: Follow point 3, the authors could try to use this method on other regions where have an extensive ground-based measurements network, for example, North America, to test their validity.

Response: Thank the referee for his/her comment. At present, the study area of our study focuses on China. In the future, we will use this method on other regions, such as North America, to test its validity.

Technical corrections:

Q2.8: Section index is wrong. For example, two “3.2” sections exist in the main text.

Response: Thank the referee for his/her comment. This issue has been fixed in the manuscript.

Q2.9: Line40: should be “van Donkelaar” instead of “Van Donkelaar”.

Response: Thank the referee for his/her comment. This name has been revised in the manuscript.

The main revision is as follows:
Hence, the approaches based on Chemical Transport Models (CTMs) (van Donkelaar et al., 2010; Wang et al., 2016; Weagle et al., 2018) or remote sensing satellites (Chen et al., 2018; Li et al., 2020; Stafoggia et al., 2019; Shtein et al., 2020; Wei et al., 2019; Yao et al., 2019; You et al., 2015) have been exploited to enlarge the spatial coverage of the PM$_{2.5}$ and PM$_{10}$ monitoring.

Q2.10: Define “DUCMASS” in the main text.

Response: Thank the referee for his/her comment. The full name of “DUCMASS” has been provided
in the manuscript.

**The main revision is as follows:**
In the meantime, the rank of DUst Column MASS density (DUCMASS) rises for the estimation of PM$_{10}$ compared to that of PM$_{2.5}$, indicating the flexibility of our approach.